DOCUMENT RESUME

ED 214 370 FL 012 816

AUTHOR	Holland, V. Melissa
TITLE	Psycholinguistic Alternatives to Readability Formulas. Document Design Project, Technical Report No. 12.
INSTITUTION	American Institutes for Research in the Behavioral Sciences, Washington, D.C.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
REPORT NO	AIR-75003-5/81-TR
PUB DATE	May 81
CONTRACT	400-78-0043
NOTE	42p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Context Clues; *Government Publications; Language Research; Language Usage; Pragmatics; *Psycholinguistics; *Readability Formulas; *Reading Comprehension; Semantics; Syntax

ABSTRACT

Features are discussed that are critical to the comprehension of texts and that readability formulas cannot handle. The critique and alternative analyses are confined to public and institutional documents and are based on research in cognitive psychology and psycholinguistics. Two types of comprehensibility complications are examined, those dealing with sentence length and syntax, and those dealing with context. It is demonstrated that readability formulas cannot grasp such variables as the effect of syntactical combinations, semantic and pragmatic relationships underlying a paragraph, and awareness of context. Examples of revisions in public documents that take account of factors beyond readability are given at the level of words, sentences, and the whole text. It is recommended that: (1) technical or special terms be tested for comprehension by the target audience; (2) content should be expressed by scenarios where possible; and (3) introductions to documents should be worded in such a way that the user will make the needed connections between items. Generally, use of design guidelines drawn from research into the comprehensibility of language and the usability of documents is recommended to writers, revisers, and evaluators of documents. (AMH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

AIR-75003-5/81-TR

DOCUMENT DESIGN PROJECT

Technical Report No. 12

ED214370

Psycholinguistic Alternatives to Readability Formulas

V. Melissa Holland

A Product of the Document Design Project funded by the National Institute of Education, Washington, D.C.
Contract No. NIE-400-78-0043

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

May 1981

FL 012 816

AMERICAN INSTITUTES FOR RESEARCH/1055 Thomas Jefferson Street, NW, Washington, DC 20007 • 202/342-5000
with Carnegie Mellon University and Siegel & Gale, Inc.

This report is part of the work being done under contract #400-78-0043 of the National Institute of Education. The project officer is Dr. Candace Miyamura of the Teaching and Learning/ Reading and Language Group. This report does not necessarily reflect the views of the funding agency.

ABOUT THE DOCUMENT DESIGN PROJECT

In September 1978, the American Institutes for Research (AIR) began the Document Design Project to foster clear and simple writing and design of public documents. The purpose of the Document Design Project (DDP) is to help make forms, regulations, brochures, and other written materials easier for people to read, to understand, and to use.

Carnegie-Mellon University and Siegel & Gale, Inc., a private firm that specializes in language simplification and language training, work with AIR on this project. Funding for the project comes from the National Institute of Education.

The Project's goal is to increase the knowledge and skills of people who produce public documents. To accomplish this goal, the Document Design Project is working on three tasks:

Task 1: To conduct theoretical and applied research studies on language comprehension, on the ways in which skilled and unskilled writers work, on problems associated with different document features.

Task 2: To work directly with government and private agencies as they produce materials for public use.

Task 3: To develop courses on writing and design for graduate students and undergraduates.

If you have questions or comments on this report or on other work of the Document Design Project, contact Dr. Janice C. Redish, Director, The Document Design Project, AIR, (202) 342-5071.

TABLE OF CONTENTS

I. Introduction

II. Critical variables beyond readability

III. Examples of revisions beyond readability

IV. Summary and Recommendations

I. Introduction

This report is taken from a paper presented at the Delaware Symposium on Language Studies, in October, 1980, which will be published in the proceedings from that conference (Di Pietro in press). The original paper was written jointly with Leo Campbell of Rutgers University. In the presentation at the Delaware conference, Mr. Campbell reviewed the history and development of readability formulas and their methodological problems. We are omitting that review here because those issues are treated extensively in previous reports from the Document Design Center.

The topic of readability, however, continues to emerge as a major issue in our work on improving the design of public documents. Explained briefly: readability is a measure of how difficult a passage of text is for an average reader. This measure is determined by a mathematical formula that counts one or two simple language variables assumed to predict ease or difficulty of comprehension. There are a number of readability formulas in use today. The variables they typically count are the average number of words per sentence and either the length of words or the frequency of the words.*

The growing use of readability formulas to evaluate educational texts, military manuals, consumer contracts, information brochures, and other materials has sparked continuing

*Word length is typically measured in terms of the number of syllables per word or the number of polysyllabic words. Word frequency is measured in terms of the presence of a word on a given list of commonly known words.

discussion and controversy. Much of this discussion has come out of our own work on improving the design of public documents. Others in the Document Design Center have treated the topic of readability on several occasions--addressing the nature, uses, and limitations of formulas as tools to measure the comprehensibility of texts (Charrow 1979; Redish 1979, 1980, 1981a).

These treatments have explored the history and development of readability formulas (Redish 1979, 1980), the mechanics and statistical rationale for selected procedures (Redish 1980), and the nature of recent variations such as the cloze test and the inclusion of specialized terms as exceptions to word frequency lists (Redish 1980, 1981a). These papers have further pointed out a number of methodological problems involved in using formulas--such as establishing grade level norms and increasing the predictive precision and intercorrelational reliability of formulas (Redish 1980, 1981a). And as a principal focus, both Redish and Charrow have demonstrated the conceptual limitations of formulas by citing examples of features obviously critical to the comprehension of texts that are ignored or contravened by readability formulas.

In the present paper, we take these limitations as our premise and move beyond listing common sense examples of the features that elude formulas. Here, we attempt to describe systematically the critical features that formulas cannot handle. We provide a basis for these features in research from cognitive psychology and psycholinguistics, and we demonstrate their

importance with illustrations from the work of our own group in analyzing, revising, and evaluating actual documents. In this way, we can begin to define valid alternatives to the formulas whose weaknesses have been well demonstrated.

More than previous treatments, we confine our critique of formulas, and our alternative analyses, to public and institutional documents rather than written materials in general. Redish (1980) describes how formulas--originally developed for educational texts used by school children--are now being extended to functional documents used by adults, such as instructions, contracts, and regulations. The present paper focuses precisely on these extensions, questions their appropriateness, and proposes alternative criteria for looking at functional documents used by varied audiences.

Both Redish and Charrow have observed that readability formulas are misused, first, by the evaluator who fails to recognize their limitations in assessing comprehensibility, and, next, by the writer who treats them as a prescription for rewriting. Both authors have recommended that alternative approaches be taken to evaluation and revision. Redish suggests that any such approach--because it must consider a range of ill-defined variables--should take the form of checklists or guidelines rather than numerical tests. It is our goal in this paper to define and demonstrate the range of variables that guidelines for effective documents must consider. (Other efforts

are underway in our group to formulate and validate specific guidelines--see American Institutes for Research, 1980).

Besides the work in our own group, we want to acknowledge the critical reviews of readability by Klare (1976), Sticht and Zapf (1976), Harris & Jakobson (1979), Kern (1979), Stevens (1980), and others (see Redish, 1980, for a review of the critics). These critics have pointed out limitations and suggested alternatives similar in broad approach to what we are presenting here in detail.

This report is organized in two sections. In Section I we demonstrate the inherent limitations of the variables used in readability formulas, drawing from studies on the comprehensibility of written language. In shaping this critique, we delineate a range of variables that go beyond readability to affect the comprehensibility of documents, and we state the basis for those variables in research from psycholinguistics and cognitive psychology. In Section II we demonstrate the practical importance of these variables in the results of diagnostic and evaluative research we have conducted on government documents. We show how these results lead to new guidelines to help designers create and evaluate comprehensible documents.

II. Critical Variables Beyond Readability

Quite aside from the methodological problems of validating, norming, and generalizing formulas, the current concept of

readability is insufficient for evaluating public documents. Even if we solve the statistical problems that surround formulas, readability cannot be a precise yardstick of how comprehensible, usable, or effective a document is. Nor can readability provide a prescription for writing and revising documents.

Why is this so? The driving assumption behind readability formulas is that sentence length and word frequency (or length) are the text features critical to predicting comprehensibility. This assumption has some support from psycholinguistic research. Several early studies showed that short sentences and familiar words are significantly easier to process than longer sentences and less frequent words. Length and frequency were viewed through verbal learning paradigms such as list recall, and through threshold paradigms, such as perception under noisy (auditory or visual) conditions (Howes & Solomon, 1952; Howes, 1957; Coleman, 1962; reviews in Cofer, 1961; Cofer & Musgrave, 1963; Dixon & Horton, 1968; Wheeler, 1970).

Recent psycholinguistic research, however, has shown that comprehensibility is considerably more complicated than the early studies indicated. Two types of complications clearly challenge readability assumptions: First are sources of comprehension difficulty in words and sentences other than their length and frequency; second are contributions of context--factors beyond the sentence level. Let us review each of these complications in Parts A and B below.

A. Other word and sentence variables

Subjects' performance in experiments with words or sentences in isolation may be a function of local factors other than length of the sentences or frequency of the words: these factors include the concreteness of words, the syntax of sentences, and a variety of less well-studied, semantic influences within words and sentences. (See Charrow 1979 for a further sample of significant word and sentence factors. See Gibson and Levin, 1975, for a comprehensive review of studies of word variables in reading.)

The concreteness of a word is usually measured by the relative speed with which it evokes a sensory image ("money" is more concrete than "resources"). Previous research (Paivio, 1969, 1971) indicates that this property is highly correlated with the frequency of a word. More important, this research shows that more concrete words are easier to understand and remember independently of length and frequency.

Similarly, a shorter sentence is more likely to be active than passive, and active sentences appear to be easier to process in a number of psycholinguistic contexts (Gough, 1965, 1966; Savin & Perchonock, 1965; Slobin, 1966). The syntactic factor is independent of length: the difference in difficulty between actives and passives is too great to be explained by differences in the number of words in the sentences.

Thus, the by-product of fitting a text to a readability formula is achieving more concrete words and simpler syntax.

These factors appear to underlie much of the measureable increase in comprehensibility of the text. Recent models of sentence processing confirm that simpler syntax accounts for more of the improved comprehension than do shorter sentences (e.g., Fodor, Bever, and Garrett, 1974).

This fact alone does not weigh against the formula, since sentence length is correlated to some degree with syntactic complexity (although it does suggest that to explain and create simpler documents we should pay attention to syntax and not to length). The complication arises when we find that simpler syntax frequently comes from lengthening sentences. This fact conflicts with formula-fitting. For example, comprehension is often found to be better when relative pronouns ("who", "which") and other words that introduce clauses ("that", "for-to") are added to sentences with embedded relative or complement clauses, as in "He felt the child trembled." vs. "He felt that the child trembled." (Fodor, Bever, & Garrett, 1974). As another example, noun strings seem to be easier when untangled into prepositional phrases, as in "kitchen table mats" vs. "mats for the kitchen table" (Charrow & Charrow, 1978; Gleitman & Gleitman, unpub.). And a relationship between two clauses is apparently better expressed explicitly, by coordination into a single long sentence ("John cried because Mary teased him."), than implicitly, by separate short sentences ("Mary teased John. He cried.") (Pearson, 1974).

The key to sentence difficulty in isolation, then, lies not in length but in syntax, as syntax interacts with people's

strategies for processing sentences (Bever, 1970; Fodor, Bever, & Garrett, 1974). Readability formulas miss this point.

What other word and sentence factors contribute to comprehensibility?

Let us start at the word level. Like concreteness, the most salient variables seem to be semantic.

First, the marked-unmarked distinction in word pairs such as "more-less," "long-short," and other polar comparisons is reflected in a comprehension and recall advantage for the unmarked ("more," "long," etc.) member of the pairs (Clark, 1969, 1973; Donaldson & Wales, 1970). Implicitly affirmative-negative pairs--"before-after," "above-below," "remember-forget"--also display systematic processing asymmetries. That is, readers understand and recall the affirmative words ("before," "above," "remember") better than their negative counterparts (Clark 1970, 1971; Just & Carpenter, 1971; Just & Clark, 1973; Carpenter & Just, 1975). Similarly, animate nouns are easier to process and recall than inanimate nouns (Rohrman, 1970). All of these differences have been explained in terms of semantic features--or, more recently, "semantic procedures"--assumed to constitute the meaning of words (Katz & Fodor, 1963; Clark, 1973; Miller & Johnson-Laird, 1976; Clark & Clark, 1977).

Again, the readability factor of frequency is highly correlated with animateness in the positive direction, and with markedness in the negative direction. Concreteness too is

correlated with animateness. But the effects cited here have been achieved with frequency, as well as concreteness, controlled. Furthermore, the semantic explanation gives a deeper and more rigorous account of these effects than does word frequency.

At the sentence level, several largely semantic variables influence comprehension:

- semantic density, or the ratio of content to function words (Perfetti, 1969a, 1969b);
- concreteness of the sentence image (distinct from the concreteness of the component words) (Paivio, 1971);
- semantic integration, or the predictability of the word combinations in the sentence (Rosenberg, 1970); and
- the semantic distance between two terms linked by the sentence (e.g., "A robin is a bird" is easier than "A chicken is a bird"), with distance given by the network or prototype configurations in models of semantic memory (Smith, Shoben, & Rips, 1974; Collins & Loftus, 1975; Rosch, 1975).

To summarize these effects, we can say that the meaning of a sentence arises from the ways in which the individual words combine. As Charrow (1979) has pointed out, it is the effect of the combinations that readability formulas cannot grasp. While formulas observe very precisely whether the constituent words are short and simple, they ignore whether the words in combination yield sense or nonsense--as in the classic anomaly, "Colorless green ideas sleep furiously."

Equally important effects of meaning come from factors beyond the word and sentence--from the domain of reference and from the surrounding discourse. Subjects in an experiment who are given an isolated sentence to understand and remember must construct a

context in which the sentence makes sense. They construct such a context only out of what they already know about situations and events. Thus meaning arises from the interaction of elements in sentences and the knowledge and circumstances of the reader (Rumelhart, 1977; Adams & Collins, 1979). A full account of comprehensibility, then, lies in the phenomena we include as context.

B. Contextual Factors

Readability formulas ignore this second set of factors--those that go beyond the word and sentence to influence comprehension. The importance of these factors is clear: Placing a word or sentence in a context changes how that word or sentence is processed. In a pioneering and by now well-known demonstration, Bransford and Johnson (1973) reported that subjects failed to comprehend a short paragraph although each word and sentence in the paragraph was simple and understandable to them. The problem was that subjects did not have access to the semantic and pragmatic relationships that underlay the paragraph and that made it coherent. Similarily, Strange and Jenkins (in Jenkins, 1977) demonstrated that the meaning of a word in a passage depends on the kind of situation referred to by the passage. Evidence like this suggests that whether words are easy or difficult--indeed the meaning of words--depends not just on their intrinsic properties, but on how they relate to each other against the larger meaning of a text or situation.

What kinds of contextual factors influence the comprehension of written language? There are a range of such factors: organization within and between paragraphs, text content and function, graphical layout, and reader characteristics. These factors are particularly important in public documents, and all are outside the grasp of current formulas.

Let us briefly consider each of these factors and cite some of the theoretical work that demonstrates their importance to language processing.

(1) Organization, the context of other sentences: The nature, variety, and explicitness of relations between sentences in a discourse--such as argument repetition, anaphoric ties, and logical connectives--have proved to be critical to comprehension.* While text linguists isolate and describe the critical levels of discourse--levels such as cohesion (Halliday & Hasan, 1976), staging or message structure (Grimes, 1975; Chafe, 1970; Halliday, 1967), and rhetorical predicates (Grimes 1975)--psychologists attempt to model how these levels enter into text processing. The seminal psychological work is on message structure (the given-new contract of Clark and Haviland, 1977), paragraph organization (the underlying argument hierarchy and the number of new arguments in a paragraph, as in Kintsch, 1974; Kintsch & van Dijk, 1978), topic accessibility and saliency in paragraphs (Kieras, 1979), and the hierarchial organization of longer discourse (Meyer, 1975). Work in computer simulation of text comprehension reinforces the

*"Discourse," as we use it here, refers to either written or spoken passages, usually more than a sentence in length.

primacy of coherence between sentences over simplicity within (Schank & Abelson, 1977; Hobbs, 1979).

On a more abstract level of organization, the overall form of the text influences comprehension. This form is specific to the text type or genre. Recent work demonstrates the psychological power of organizational forms, called schemas or grammars, in the processing of fairly stereotyped kinds of discourse like stories and research articles (Rumelhart, 1975; Mandler & Johnson, 1977; Kintsch & van Dijk, 1978; Stein & Glenn, 1979). Public documents such as regulations, applications, and contracts, characteristically conform to prototypical forms of organization. However, the potency of these prototypes in people's use of documents remains largely unexplored (Holland & Redish, in press).

(2) Content, the context of reference and knowledge: We referred earlier to significant semantic variables within words and sentences. A fuller account of meaning is given by the reference a sentence or passage builds to a situation or set of relationships, and the contact it makes with a reader's framework of knowledge. This knowledge appears to take the form of highly organized cognitive frames, scripts, or more generally, schemas. When a text engages this knowledge, the reader can use it to create a referential context that will "account for" the text (Anderson, 1977, 1978; Rumelhart & Ortony, 1977; Schank & Abelson, 1977; Adams & Collins, 1979; Tannen, 1979; Bransford, 1979). When the reader is unable to invoke a schema (or schemas) for the text, comprehension suffers (Bransford & Johnson, 1973; Dooling &

Mullet, 1973). This may be because the reader does not possess the appropriate schema, or because the wording and structure of the text are ineffective in tapping it.

Furthermore, the way a passage is interpreted and recalled depends on which of a number of possible schemas the reader brings to bear on that passage (Anderson, Reynolds, Schallert, & Goetz, 1977; Pichert & Anderson, 1977; Spiro, 1977; Anderson & Pichert, 1978). This in turn depends on the particular knowledge background of readers (Anderson et al., 1977), as well as on their immediate orientation and instructional set (Pichert & Anderson, 1977).

(3) Function, the context of tasks: The functions of a text, and the corresponding goals of the reader in approaching it, can vary. Public documents, in particular, serve a range of real-world uses. To illustrate: a document can be for learning and remembering (a brochure on the warning signs of cancer), for answering selected questions (a regulation or a warranty), for giving instructions (a recipe or a how-to kit), for gathering information (an application form), or for informing immediate decisions (a contract, a product label).

The prominence of function in predicting text difficulty emerges clearly in recent studies. For example, the simplest format for presenting rules, regulations, and other information--whether by prose or alternative formats such as flowcharts--varies with the purposes of the information--e.g., whether you are to learn it or use it for reference (Wright &

Reid, 1973; Kern, Sticht, Welty, & Hauke, 1976). Similarly, the clearest statement of prose headings and whole passages in a regulation (whether by definitions or examples), seems to depend on the reader's needs in approaching the regulation (Flower, Hayes, & Swarts, 1980; Swarts, Flowers, & Hayes, 1980). To predict the functional difficulty of a document with reasonable accuracy, we need a detailed task analysis in terms of the cognitive and physical demands the document places on the reader (see also Spiro & Taylor, 1980).

(4) Graphics, the context of the physical document: Studies reviewed by Wright (1978) and Macdonald-Ross (1977) demonstrate the importance of format to comprehension. Work by Wright and others (Wright & Reid, 1973; Wright, 1977; Landa, 1976; Follettie, 1978, 1979), for example, suggests that tabular or algorithmic presentations are superior to prose for particular kinds of content. Headings, type face, color, indentation, blocking, and white space are other graphics variables that affect the ease of understanding and using documents and other kinds of texts (e.g., Coleman & Kim, 1961; Tinker, 1963; Ausubel, 1978; Frase & Schwartz, 1979).

(5) Reader characteristics, the context of users: Who the readers are is important for predicting how difficult a word, a sentence, or the document as a whole will be (Cole & Scribner, 1974; Rodriguez, 1980). Readers' culture-based expectations and abilities (beyond the reading grade level which formulas consider) are the characteristics that appear crucial. Ethnographic

research in particular is beginning to show that different populations--the elderly or Hispanics in this country, for example--have different ways of approaching and understanding text, have more or less experience with specific types of text, and have culture-specific meanings for some standard words and phrases.

Thus, over and above readability variables, a range of interacting factors influence how well a document will be understood. We have not attempted to list all possible influential factors, but instead to give an idea of the scope and variety of principles that complicate the picture given by readability formulas. While sentence length and word frequency do contribute to the difficulty of a document, a number of equally important variables elude and sometimes run at cross purposes to the formulas (as with the syntactic features illustrated earlier).

Reliance on readability formulas in revising and evaluating documents obscures the importance of the other critical variables. Formula-fitting may lull writers into a false sense of having solved all the communication problems of a document, leading them to ignore aspects they might otherwise consider. The accumulated evidence suggests to us that readability level can appropriately serve only as a filter, to decide which texts are <u>not</u> acceptable. It is not appropriate as a ruler of effectiveness or as a design goal.

The goals of writing and revising should instead consider several levels of interacting factors--document function, audience, expectations, organization, coherence, syntax, and layout, for example. Yet we know little about how these factors work separately and less about how they work together; and we can assume their interactions will vary over the gamut of structures, lengths, and cognitive tasks required by documents. The principles for applying these factors do not permit algorithmic formulation, but better take the shape of guidelines that can extend flexibly to different documents (cf. Kern, 1979; Redish, 1979).

There are sets of such guidelines in use by document designers now (e.g., Morris, 1972; Post & Price, 1974; Kern, et. al, 1976; Redish, Charrow, Felker, & Sheard, 1979; American Institutes for Research, 1980; Redish, 1981b; Redish, Landesman, Peck, & Felker, 1981). When these guidelines do not provide answers for specific problems, and these problems are pressing enough, the designer must consult the cognitive, psycholinguistic, and human factors literature, or conduct ad hoc research with the document at hand and its intended audience.

III. Examples of Revisions Beyond Readability

In the remaining part of the paper, we will describe some revisions in public documents that do take account of factors beyond readability. We have selected three examples--one at the level of words, one at the sentence level, and one at the whole

text level. These examples illustrate four of the factors we have mentioned: culture-specific reader characteristics, word meaning, sentence meaning vis a vis cognitive frames, and organizational or text schemas. All revisions demonstrably improved comprehension and performance, and all are changes which a formula cannot capture.*

These revisions came out of practical research stimulated by a need to solve problems not currently handled by published guidelines. The results of this research have added to our understanding of the variables relevant to comprehension and have allowed us to develop new guidelines for creating and judging effective documents.

The first example of a revision outside of readability comes from a study of the current District of Columbia Medicaid Recertification Form. Working through the local Medical Assistance Office, we observed problems that recipients have in understanding and using this form and then designed a revision to alleviate the problems (Charrow, Holland, Peck, & Shelton, 1980).

One kind of change we made was simply at the word level, to take account of how special reader characteristices affect word meaning. The people who have to use the Medicaid form are poor, not well-educated, often old, and frequently speakers of a

*Some of these examples came from application forms, with fill-in-the-blank sections to which readability scores do not apply (since scores are computable only on text consisting of complete sentences). Bureaucratic writers have nevertheless attempted to extend readability principles to forms (shortening sentences and words), as the sole guide for designing and revising. Part of our argument is against these extensions.

non-standard dialect, or not native speakers of English. Our observations showed that people who fill out the form have a good deal of trouble with terminology. This was predicatable. But sometimes the nature of the trouble surprised us. For example, many users misinterpreted the term "family member," which recurs in important questions throughout the form. Or, we should say, they misinterpreted the term relative to the dialect of the document designer.

The designer intends the phrase to refer to people in the immediate, nuclear family of the Medicaid recipient, all covered by a common Medicaid card. This meaning is not spelled out, but is implied by listing the names of the people included and placing them in a box labeled "family members" near the top of the form. The designer assumes the meaning is clear: it is defined by the graphics, and the phrase itself is simple, familiar, and has a unique meaning to the designer.

But people in the Medicaid population typically have extended families--a network of kin, often reared together, and all considered immediate family. As might be expected, many respondents answered questions about family employment and resources as if those questions referred to nephews, cousins, sisters-in-law, and others not covered by the recipient's Medicaid card. The list in the box was not salient enough to overcome this more natural definition of "family members." Our solution, where space permitted early in the form, was to spell out the reference

("the people in the box . . .") instead of relying on the repeated label ("family members").

Clearly, we would never have made this kind of change without some observations of Medicaid clients actually using the form. A readability formula tells generally whether a word is likely to be in the vocabulary of a reader, but it does not tell about meaning. Particularly, it doesn't tell whether there are unusual meanings attached to a word, meanings specific to a document's target audience. While substituting high- for low-frequency words is a good idea, and we did it often in revising the Medicaid form, we found it necessary to go beyond this principle, paying attention to meaning and use in the intended audience.

What sort of guideline can we develop based on results like these? We would recommend that all technical or special terms be tested for understanding by the target audience. If a term is not understood, or is understood differently from the designer's intention, then the term should be changed or defined explicitly.

In the next example, we move to the sentence level and consider the factor of content: specifically, the relation between the meaning of a sentence and the reader's frame of reference. The revision we will look at was developed by Flower and Hayes at Carnegie-Mellon University (see Flower, Hayes, & Swarts, 1980). The Carnegie-Mellon group studied how people make sense of government regulations. They tape-recorded people reading regulations aloud and trying to interpret them. Here is what the reading protocols showed: to make the information in a

passage meaningful, readers tended to turn definitions and instructions into a small dramatized "scenario"--a scene in which someone does something.

Take, for example, a paragraph headed "Applying for Bank Loans" and starting with the sentence, "The procedure to apply for bank loans consists of . . ." A reader might turn this sentence into, "Now say if I wanted to apply for a bank loan, then I would . . .", or "Now what if this fellow wanted to apply for a bank loan, then he . . ."

In trying to understand what a regulation says, readers relied on the scenario principle. They took abstract definitions, with nominalizations like "procedure" as subject, and expressed them in the form of story-like events. These events had human agents as subjects, doing an action in the context of a time and place. The event frame makes information concrete, imagible, and familiar through relationships like agent-object and action-goal. As Flower, et al., noted, the scenario seems to be a basic structure of experience and a way we naturally think.

When the authors rewrote regulations according to the scenario principle, and evaluated them with comprehension tests, they found that the revised regulations were easier to understand and use. Flower, et al., explained their findings in terms of gearing the document to the goals and needs of the reader. People come to regulations not to learn definitions, but to act, to make decisions, to answer questions. The human-centered scenario aids actions and decisions better than do definitions.

Interestingly, Flower, et al., found that their readers spent little time wrestling with hard words or trying to reduce long sentences. In fact, creating scenarios usually meant making sentences longer--counter to the predictions of readability formulas. Of course, the scenario does use short, simple words, often to replace longer words--so in most cases it does not violate readability assumptions. The important point is that the scenario principle more accurately predicts and better explains the effects of these changes.

Independently of the research of Flower, Hayes, and Swarts, the work done on the Medicaid form at the American Institutes for Research revealed the need for something like a scenario principle to improve people's understanding of that document. Our problem analysis included watching Medicaid caseworkers helping clients work through their recertification forms. We paid particular attention to the "translations" caseworkers made of the document in an effort to clarify it for the client. A scenario-like translation was the preferred strategy to explain complex concepts like "support," "cash on hand," "continued eligibility" and even "recertification." For example, for one of the items on the current form, "amount of support from friends or relatives: _____," caseworkers commonly said something like, "Does a friend or relative give you money, food, or a place to live?" We took our cue from the caseworkers. In the revised form, we itemized, turned labels into questions, and used brief scenario-like translations for sentences with abstract or general terms, where space permitted.

What guideline can we draw from these results? Very simply: express content by scenarios where possible.

We are not suggesting that this is the way to go with all documents. People who deal frequently with bureaucratic forms are familiar with terms like "support" and "recertification" and are in the habit of reading labels with blanks (e.g., "support from relatives: _____") as instructions to respond. Rather, we are suggesting that the kind of revisions one makes in a document should stem from knowing the audience. One way this knowledge can be gained is through ad hoc observations of people in that audience using the document in question (see Rose, 1981, who recommends a similar approach to revising documents).

Moreover, we are aware that the best clarifications cannot always be incorporated in a revision. Space constraints may prohibit the consistent use of translations like the ones we attempted in the Medicaid form. But if readers can at least start with explicit questions or imperatives, as well as concrete scenarios and examples, then they appear to proceed through the rest of the form more easily.

In the next example, we move to the whole text level and look at the factors of organizational schemas and readers' assumptions about text. Up to now, we have considered ways to facilitate comprehension and use that pay attention to words, sentences, and larger units within a document. Such revisions could be viewed as reducing the reader's processing load--the load of recognizing words, breaking text into manageable-size chunks, holding and

storing them in memory, translating concepts into personal terms, drawing inferences, and generally extracting meaning and the conditions for action.

But beyond looking at the processing demands imposed by a document, it may help the writer/reviser to consider what assumptions lead the user to take on those demands in the first place. Why does the user attempt to find meaning, draw inferences, and make connections among items across the form? We want to argue that one assumption the user makes--in fact, a prerequisite for understanding--is that the document is discourse. This means seeing the document as reflecting a coherent intention on the part of the designer, seeing it as purposive and communicative.

Indeed, documents are coherent. Even forms or contracts--which consist of numbered and seemingly unconnected items--are more than a concatenation of these items: they display selected characteristics of more traditional, continuous prose. These characteristics include cohesiveness in the form of anaphora, word repetition, and other kinds of ties between items (see Halliday & Hasan, 1976); given-new conventions in the presentation of ordered information (see Clark & Haviland, 1977); inferred bridges between items as required by the given-new convention; and an assumption of cooperation shared between reader and writer (see Grice, 1975). (How forms manifest these and other characteristics is discussed in Charrow, et al., 1980 and Holland, 1980.) Readers must tap these characteristics if they are going

to understand the document. And to tap the discourse conventions, readers must see the document as purposive communication.

That documents are communication may seem obvious enough. We have evidence, though, that some readers do not see documents this way. Many recipients of the Medicaid form used in the AIR case study indicated that they did not understand what that form was about, why they were getting it, or even what it was generically. The problem is, these people do not normally deal with documents, do not see them as a type of text, and have no framework for classifying or understanding them. The introduction to the form (a few lines at the top of the first page), which is supposed to state the purpose, was apparently not seen as intending to explain anything at all. Further, each item on the form tended to be treated as new, independent, and unconnected to previous items. This treatment led to predictable patterns of errors (Holland, 1980; Holland & Redish, in press).

A second example of readers failing to recognize and use discourse phenomena to help them get through a form comes from protocol analysis of people filling out the government's job application form (SF-171). Preliminary analyses of the protocols have been performed by Dixie Goswami of the Document Design Center, and are reported in Holland and Redish (in press). Results suggest that people who have difficulties with the form and who fail to represent their personal, job-relevant data appropriately are just those people who fail to check for connections between items and whose protocol comments are

singularly item-bound. By contrast, the more expert form-users refer to the rhetorical context of the document, acknowledge the purposes and perspectives of the agency administering the form, and call on that knowledge as a guide to presenting themselves on the form.

We feel that, when the readers in both these examples ignored connections and purpose, it was not because they were unable to handle text, but because they failed to perceive the form as text.

This was our solution in the case of the Medicaid form: we re-wrote the introduction as a letter. The revised form began "Dear (name of client)," and continued with the introductory statements on succeeding lines. We reasoned that this letter "schema" would be familiar to readers, would signal communication and purpose, and would lead them to seek coherence.

Although we do not yet have sufficient data to judge the effectiveness of rewriting the introduction as a the letter for this population, the design gives early indications of helping people better handle the Medicaid form. This is the kind of global organizational principle which, again, goes beyond what readability formulas can tell us about how to help people deal with documents.

IV. Summary and Recommendations

Even if readability formulas were correctly validated using educational materials and school age populations, their validity would not necessarily extend to public documents. We have shown that the principles underlying formulas are often in conflict with a document designer's specific goals. But we have tried not to dismiss formulas entirely.

For the writer and editor, the readability variables of sentence length and word frequency could be included in the range of criteria--syntactic, semantic, organizational, graphic, cultural--for constructing an effective document. For the evaluator, formulas could serve as a rough filter, to precede a more specific and meaningful test of effectiveness.

Formulas can also be used to indicate change in the drafting and revising process. But as a predictor of comprehensibility, formulas simply don't work very well, especially for text in which the reader/user must process diffuse types of information and perform any number of tasks with the material. Public use documents are clearly text of that kind.

As an alternative to formulas for guiding writers and revisers and for evaluating documents, we recommend the use of design guidelines that draw from research into the comprehensibility of language and the usability of documents. In this paper, we defined the nature and scope of the variables relevant to guidelines, and we presented samples of our own

ongoing research into the effects of these variables in documents. The results of this research are helping shape the set of design guidelines our group is currently producing and validating.

REFERENCES

Adams, M.J. & Collins, A. A schema-theoretic view of reading. In R.O. Freedle (Ed.), New directions in discourse processing, Vol. 2. Norwood, N.J.: Ablex, 1979.

American Institutes for Research. Guidelines for document designers: A working draft. (Document Design Project Technical Report 75002-IR-4/80) Washington, D.C.: American Institutes for Research, April 1980.

Anderson, R.C. The notion of schemata and the educational enterprise: General discussion of the conference. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1977.

Anderson, R.C. Schema-directed processes in language comprehension. In A.M. Lesgold, J.W. Pellegrino, S.D. Fokkema, & R. Glaser (Eds.), Cognitive psychology and instruction. New York, N.Y.: Plenum Press, 1978.

Anderson, R.C. & Pichert, J.W. Recall of previously unrecallable information following a shift in perspective. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 1-12.

Anderson, R.C., Reynolds, R.E., Schallert, D.L. & Goertz, E.T. Frameworks for comprehending discourse. American Educational Research Journal, 1977, 14, 367-381.

Ausubel, D. In defense of advanced organizers: A reply to the critics. Review of Educational Research, 1978, 48, 251-257.

Bever, T.G. The cognitive basis for linguistic structures. In J.R. Hayes (Ed.), Cognition and the development of language. New York, N.Y.: Wiley, 1970.

Bransford, J.D. Human cognition: Learning, understanding, and remembering. Belmont, Calif.: Wadsworth Press, 1979.

Bransford, J.D. & Johnson, M.K. Considerations of some problems of comprehension. In W.G. Chase (Ed.), Visual information processing. New York, N.Y.: Academic, 1973.

Carpenter, P. & Just, M.A. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 1975, 82, 45-73.

Chafe, W. Meaning and the structure of language. Chicago, Ill.: University of Chicago Press, 1970.

Charrow, V.R. Let the rewriter beware. Washington, D.C.: American Institutes for Research, 1979.

Charrow, V.R. & Charrow, R.P. The comprehension of standard jury instructions: A psycholinguistic approach. Working paper, Center for Applied Linguistics, 1978.

Charrow, V.R., Holland, V.M., Peck, D.G., & Shelton, L.V. Revising a Medicaid recertification form: A case study in the document design process. Washington, D.C.: American Institutes for Research, 1980.

Clark, H.H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404.

Clark, H.H. How we understand negation. Paper presented at COBRE Workshop on Cognitive Organization and Psychological Processes. Huntington Beach, Calif.: August, 1970.

Clark, H.H. The chronometric study of meaning components. Paper presented at the CRNS Colloque International sur les Problemes Actuels de Psycholinguistique. Paris, France: December, 1971.

Clark, H.H. Semantics and comprehension. In T.A. Sebeok (Ed.), Current trends in linguistics (Vol. 12). The Hague: Mouton, 1973.

Clark, H.H. & Clark, E.V. Psychology and language. New York, N.Y.: Harcourt, Brace & Jovanovich, 1977.

Clark, H.H. & Haviland, S.E. Comprehension and the given-new contract. In R.O. Freedle (Ed.), Discourse production and comprehension, Vol. I. Norwood, N.J.: Ablex, 1977.

Cofer, C.N. (Ed.). Verbal learning and verbal behavior. New York, N.Y.: McGraw-Hill, 1961.

Cofer, C.N. & Musgrave, B.S. (Eds.). Verbal behavior and learning: Problems and processes. New York, N.Y.: McGraw-Hill, 1963.

Cole, M. & Scribner, S. Culture and thought. New York, N.Y.: Wiley, 1974.

Coleman, E.B. Improving comprehensibility by shortening sentences. Journal of Applied Psychology, 1962, 46, 131-134.

Coleman, E.B. & Kim, I. Comparison of several types of typography in English. Journal of Applied Psychology, 1961, 45, 1252-1267.

Collins, A. & Loftus, E.F. A spreading activation theory of semantic processing. Psychological Review, 1975, 82, 407-428.

Di Pietro, R. (Ed.) Linguistics and the professions. Norwood, N.J.: Ablex, in press.

Dixon, T.R. & Horton, D.L. (Eds.) Verbal behavior and general behavior theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Donaldson, M. & Wales, R. On the acquisition of some relational terms. In J.R. Hayes (Ed.), Cognition and the development of language. New York, N.Y.: Wiley, 1970.

Dooling, D.J. & Mullet, R.L. Locus of thematic effects in retention of prose. Journal of Experimental Psychology, 1973, 97, 404-406.

Flower, L.S., Hayes, J.R., & Swarts, H. Revising functional documents: The scenario principle. (Document Design Project Technical Report No. 10). Pittsburgh, Penn.: Carnegie-Mellon University, March 1980.

Fodor, J.A., Bever, T.G., & Garrett, M.F. The psychology of language. New York, N.Y.: McGraw-Hill, 1974.

Follettie, J.F. Numerical display research: Review and discussion. (Technical Report 2-78-01). Los Alamitos, Calif.: SWRL Educational Research and Development, 1978.

Follettie, J.F. Graph-using profiencies of upper elementary schoolers. (Technical Report 63). Los Alamitos, Calif.: SWRL Educational Research and Development, 1979.

Frase, L. & Schwartz, B. Typographical cues that facilitate comprehension. Journal of Educational Psychology, 1979, 71, 197-206.

Gibson, E.J. & Levin, H. The psychology of reading. Cambridge, Mass.: MIT Press, 1975.

Gleitman, L.R. & Gleitman, H. Phrase and paraphrase. University of Pennsylvania, unpublished manuscript.

Gough, P.B. Grammatical transformations and speed of understanding. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 107-111.

Gough, P.B. The verification of sentences: The effects of delay of evidence and sentence length. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 492-496.

Grice, H.P. Logic and conversation. William James Lectures, Harvard University, 1967, unpublished. Lecture 2 published in P. Cole & J.L. Morgan (Eds.)., Studies in syntax, III. New York, N.Y.: Seminar Press, 1975.

Grimes, J.E. The thread of discourse. The Hague: Mouton, 1975.

Halliday, M.A.K. Notes on transitivity and theme in English: II. Journal of Linguistics, 1967, 3, 199-244.

Halliday, M.A.K. & Hasan, R. Cohesion in English. London, England: Longman, 1976.

Harris, A.J. & Jacobson, M.D. A framework for readability research: Moving beyond Herbert Spencer. Journal of Reading, 22, 5, 1979, 390-398.

Hobbs, J. Coherence and coreference. Cognitive Science, 1979, 3, 67-90.

Holland, V.M. Revising a government document: The case of the Medicaid Recertification Form. Paper presented at the Symposium on Writing and Designing Documents: Research and Practical Solutions. Carnegie-Mellon University, October, 1980.

Holland, V.M. & Redish, J.C. Strategies for understanding forms and other public documents. In D. Tannen (Ed.), Proceedings of the Georgetown Round Table on Discourse: Text and Talk, Washington, D.C., in press.

Howes, D. On the relation between intelligibility and frequency of occurrence of English words. Journal of the Acoustic Society of America, 1957, 29, 296-305.

Howes, D. & Solomon, R.L. Visual duration thresholds as a function of word probability. Journal of Experimental Psychology, 1952, 44, 260-266.

Jenkins, J.J. Context conditions meaning. Paper presented at the Midwestern Psychological Association. Chicago, Ill.: May, 1977.

Just, M.A. & Carpenter, P.A. Comprehension of negation with qualification. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 244-253.

Just, M.A. & Clark, H.H. Drawing inferences from the presuppositions and implications of affirmative and negative sentences. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 21-31.

Katz, J. & Fodor, J. The structure of a semantic theory. Language, 1963, 39, 170-220.

Kern, R.P. Usefulness of readability formulas for achieving Army readability objectives: Research and state-of-the-art applied to the Army's problem. Fort Benjamin Harrison, Ind.: Technical Advisory Service, U.S. Army Research Institute, 1979.

Kern, R.P., Sticht, T.G., Welty, D. & Hauke, R.N. Guidebook for the development of army training literature. Washington, D.C.: U.S. Army Research Institute for the Behavioral and Social Sciences, HumRRO, November 1976.

Kieras, D.E. The role of global topics and sentence topics in the construction of passage macrostructure. Technical Report No. 6. Tucson, Ariz.: University of Arizona, 1979.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W. & van Dijk, T.A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Klare, G.R. A second look at the validity of readability formulas. Journal of Reading Behavior, 1976, 8, 129-152.

Landa, L. Instructional regulation and control: Cybernetics, algorithmization, and heuristics in education. Englewood Cliffs, N.J.: Educational Technology, 1976.

Macdonald-Ross, M. Graphics in texts. In L.S. Shulman (Ed.), Review of research in education. Itasca, Ill: Peacock, 1977.

Mandler, J.M. & Johnson, N.S. Remembrance of things parsed. Story structure and recall. Journal of Verbal Learning and Verbal Behavior, 1977, 9, 111-151.

Meyer, B.J.F. The organization of prose and its effect on recall. Amsterdam, The Netherlands: North-Holland, 1975.

Miller, G.A. & Johnson-Laird, P.N. Language and perception. Cambridge, Mass.: Harvard University Press, 1976.

Morris, J.O. Make yourself clear! New York, N.Y.: McGraw-Hill, 1972.

Paivio, A. Mental imagery in associative learning and memory. Psychological Review, 1969, 76, 241-263.

Paivio, A. Imagery and verbal processes. New York, N.Y.: Holt, 1971.

Pearson, P.D. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. Reading Research Quarterly, X (2), 1974-75, 155-192.

Perfetti, C. Lexical density and phrase structure depth as variables in sentence retention. Journal of Verbal Learning and Verbal Behavior, 1969a, 8, 719-724.

Perfetti, C. Sentence retention and the depth hypothesis. Journal of Verbal Learning and Verbal Behavior, 1969b, 8, 101-104.

Pichert, J.W. & Anderson, R.C. Taking different perspectives on a story. Journal of Educational Psychology, 1977, 69, 309-315.

Post, T. & Price, H. Requirements and criteria for improving reading comprehension of technical manuals. (No. 0024-74-C-5426). Falls Church, Va.: Biotechnology, 1974.

Redish, J.C. Readability. Washington, D.C.: American Institutes for Research, 1979.

Redish, J.C. Readability. In D. Felker (Ed.), Document design: A review of the relevant research. Washington, D.C.: American Institutes for Research, 1980.

Redish, J.C. Understanding the limitations of readability formulas. IEEE Transactions on Professional Communication. 1981(a), 24(1), 46-48.

Redish, J.C. How to write regulations (and other legal documents) in clear English. In R.A. Givens (Chairman), Drafting documents in plain language-1981. New York: Practising Law Institute, 1981(b).

Redish, J.C., Charrow, V., Felker, D. & Sheard, V. Simplifying documents - A workshop manual. Washington, D.C.: American Institutes for Research, 1979.

Redish, J.C., Landesman, J., Peck, D. & Felker, D. Designing simpler forms - A workshop manual. Washington, D.C.: American Institutes for Research, 1981.

Rodriguez, B.R. Document considerations for Spanish-speaking users. Paper presented at the annual meeting of the American Educational Research Association, Boston, Mass., 1980.

Rohrman, N.L. More on the recall of nominalizations. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 534-536.

Rosch, E.H. Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 1975, 104, 192-233.

Rose, A.M. Identifying problems in critical public documents. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, Calif., 1981.

Rosenberg, S. The recall of verbal material accompanying semantically well-integrated and semantically poorly-integrated sentences. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 534-536.

Rumelhart, D.E. Notes on a schema for stories. In D. Bobrow and A. Collins (Eds.), Representation and understanding: Studies in cognitive science. New York, N.Y.: Academic Press, 1975.

Rumelhart, D.E. Toward an interactive model of reading. In S. Dornic (Ed.), Attention and performance VI. Hillsdale, N.J.: Erlbaum, 1977.

Rumelhart, D.E. & Ortony, A. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro, and W.E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1977.

Savin, H.B. & Perchonock, E. Grammatical structure and the immediate recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 348-353.

Schank, R.C. & Abelson, R.P. Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. Hillsdale, N.J.: Erlbaum, 1977.

Slobin, D.I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-277.

Smith, E.E., Shoben, E.J. & Rips, L.J. Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review, 1974, 81, 214-241.

Spiro, R.J. Remembering information from text: The "state of schema" approach. In R.C. Anderson, R.J. Spiro, and W.E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1977.

Spiro, R. J. & Taylor, B.M. On investigating children's transition from narrative to expository discourse: The multidimensional nature of psychological test classification. (TR-194). University of Illinois: Center for the Study of Reading, December 1980.

Stein, N.L. & Glenn, C.G. An analysis of story comprehension in elementary school children. In R.O. Freedle (Ed.), New directions in discourse processing, Vol. II. Norwood, N.J.: Ablex, 1979.

Stevens, K.C. Readability formulae and McCall-Crabbs Standard Test Lessons in Reading. The Reading Teacher, 33, 4, 1980, 413-415.

Sticht, T.G. & Zapf, D.W. (Eds.) Reading and readability in the armed services. (HumRRO Technical Report 76-A) Alexandria, Va.: HumRRo, 1976.

Swarts, H., Flower, L.S. & Hayes, J.R. How headings in documents can mislead readers. (Document Design Project Technical Report No. 9). Pittsburgh, Penn.: Carnegie-Mellon University, April 1980.

Tannen, D. What's in a frame? Surface evidence for underlying expectations. In R.O. Freedle (Ed.), New directions in discourse processing, Vol. II. Norwood, N.J.: Ablex, 1979.

Tinker, M. Legibility of print. Ames, Iowa: Iowa State University, 1963.

Wheeler, D.D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.

Wright, P. Decision-making as a factor in the ease of using numerical tables. Ergonomics, 1977, 20, 91-96.

Wright, P. Feeding the information eaters: Suggestions for integrating pure and applied research on language comprehension. Instructional Science, 1978, 7, 249-312.

Wright, P. & Reid, F. Written information: Some alternatives to prose for expressing the outcomes of complex contingencies. Journal of Applied Psychology, 1973, 57, 160-166.